

FEATURES	BENEFITS
<ul style="list-style-type: none"> <li>• Simple overview of production statistics</li> <li>• Filter date, time, operator, and SN</li> <li>• Automatic data collection</li> <li>• Open output format (CSV)</li> <li>• Statistics of verdicts and single values</li> <li>• Use for regular statistics</li> <li>• Observe distribution and time course of single values</li> </ul>	<ul style="list-style-type: none"> <li>• Supervision of production</li> <li>• Overview of verdicts and single value results</li> <li>• Simple interface to 3<sup>rd</sup> party software</li> <li>• Generate customized reports of yield and single values</li> <li>• Keep track and identify drifts of single values (e.g. coil shift versus time)</li> <li>• Generate customized reports</li> </ul>

This complementary module delivers a quick statistical overview on test results of production lines. Log files, generated by *Klippel QC* software are processed and the results are shown as *Verdict statistics* and *Single value statistics*.

The data can be filtered by date/time, operator, SN and single value names. Different production lines can be parsed, by copying the log files to a network folder that is accessible for the statistic module.

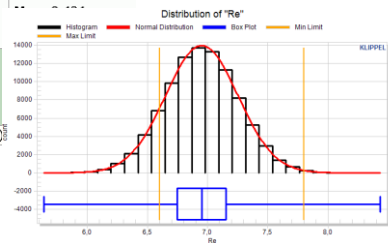
Third party statistics software can be used by exporting a CSV file of the parsed data.

#### Verdict statistics

	Overall	Resp	Level	Pol
<b>Yield (%)</b>	<b>64.88</b>	<b>97.01</b>	<b>78.72</b>	<b>97.03</b>
<b>Valid</b>	94206	94206	94206	94206
<b>Pass</b>	61119	91389	74156	91409
<b>Warning</b>	0	0	0	0
<b>Fail</b>	33087	2817	20050	2797

#### Single value statistics

	Level	Re
<b>All</b>	x: 112.8 Min: 102.8 Max: 123.1 σ: 2.404 N: 94206	x: 6.951 Min: 5.639 ..
<b>Passed</b>	x: 112.9 Min: 110 Max: 116 σ: 1.56 N: 61119	
<b>Limits</b>	[110 .. 116] (94206)	



#### Application:

- Overview of yield
- Regular (daily, weekly, ...) statistics
- Overview of verdicts
- Overview of single value results

#### CONTENT:

1	Summary .....	2
2	Definition of Statistics .....	2
3	Input Parameters.....	4
4	Placeholder Tokens for Date/Time Input .....	5
5	Handling Large Amounts of Data .....	6

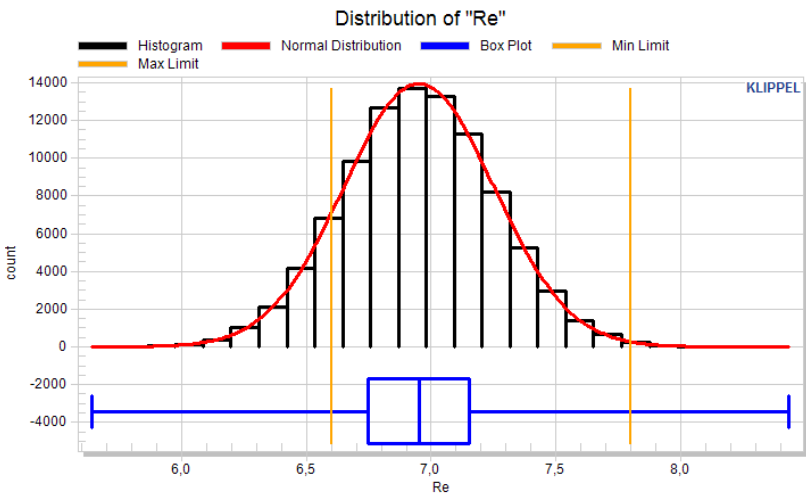
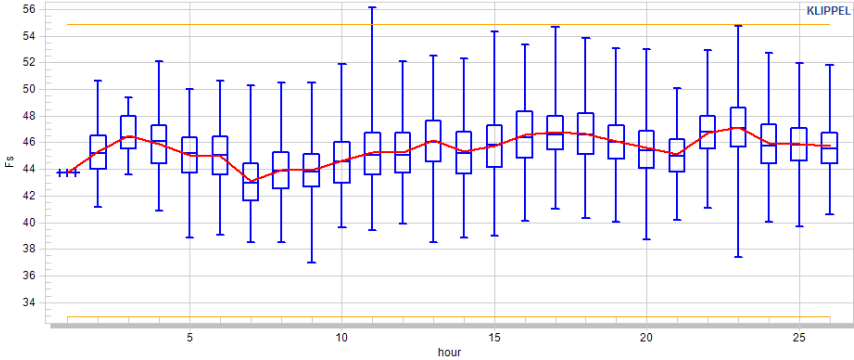
## 1 Summary

The *Yield Statistics* module processes log files produced by the *Klippel QC* software. No hardware is required for executing this module.

The *Yield Statistics* can be used to generate regular statistics. Refer to application note *AN 48 Yield Statistics* for examples.

## 2 Definition of Statistics

<b>Valid Measurements</b>	<p>The verdict counters are fed with data from the parsed log files. The following counters are available and represent valid measurements:</p> <ul style="list-style-type: none"> <li>• Pass</li> <li>• Warning</li> <li>• Fail</li> </ul> <p>The total number of valid measurements is shown additional to these counters. Please note that the verdict Warning equals a Pass from the yield point of view.</p>
<b>Yield</b>	<p>The yield <math>Y</math> is defined as the ratio of number of passed DUTs <math>N_{\text{pass}}</math> (<i>Pass</i> and <i>Warning</i>) to total number of all valid measurements <math>N_{\text{total}}</math> (<i>Pass</i>, <i>Warning</i> and <i>Fail</i>) in percent</p> $Y = \frac{N_{\text{pass}}}{N_{\text{total}}} \cdot 100\%.$ <p>The yield is calculated for all available individual verdicts. The overall yield is highlighted.</p>
<b>Invalid Measurements</b>	<p>Invalid measurements (verdicts <i>Invalid</i>, <i>Noise</i> and <i>Void</i>) do not contribute to the calculation of the yield. The invalid measurements are shown below the valid measurements.</p> <p>Please note, that the invalid measurement verdicts are not shown if they are all zero.</p>
<b>SINGLE VALUE STATISTICS</b>	
<b>All / Pass</b>	<p>Each of the following statistical values is calculated both for the overall passed single values and for all valid single values.</p>
<b>Mean (<math>\bar{x}</math>)</b>	<p>The mean <math>\bar{x}</math> of a single value is defined as the ratio of the sum of the single value <math>\sum_{n=1}^N x(n)</math> to the number of samples <math>N</math></p> $\bar{x} = \frac{\sum_{n=1}^N x(n)}{N}.$ <p><b>Note:</b> The responsibility to interpret the single value statistics lies with the user. For example: a linear averaged level differs from terms like “averaged SPL”. The total number of samples equals the number of valid measurements (verdicts Pass, Warning and Fail).</p>
<b>Min/Max</b>	<p>The smallest/largest element found in the parsed data</p> $x_{\min} = \min_n x(n),$ $x_{\max} = \max_n x(n).$
<b>Deviation (<math>\sigma</math>)</b>	<p>The standard deviation <math>\sigma</math> of the single value is defined as follows:</p>

	$\sigma = \sqrt{\left(\sum_{n=1}^N x(n)^2\right) - \frac{1}{N}\left(\sum_{n=1}^N x(n)\right)^2}$ $= \sqrt{\left(\sum_{n=1}^N x(n)^2\right) - N \cdot \bar{x}^2}$
<p><b>Samples (N)</b></p>	<p>The number of valid measurements due to the overall verdict.</p>
<p><b>Limits</b></p>	<p>Minimum and maximum value of the limit with absolute number of occurrences of this combination of minimum and maximum limit. If the limits change (e.g. in different log files), the different limits are listed one below the other. If no limit is available for a single value “-” is shown.</p>
<p><b>DISTRIBUTION AND TIME COURSE ANALYSIS</b></p>	
<p><b>Histogram</b></p>	<p>Distribution analysis of single values as a histogram plot:</p>  <p>The figure shows a histogram of 'Re' values. The x-axis is labeled 'Re' and ranges from 6.0 to 8.0. The y-axis is labeled 'count' and ranges from -4000 to 14000. A red normal distribution curve is overlaid on the histogram. A blue box plot is shown below the histogram. Two vertical orange lines represent the 'Min Limit' and 'Max Limit' at approximately 6.6 and 7.8 respectively. The legend indicates: Histogram (black bars), Normal Distribution (red line), Box Plot (blue box), Min Limit (orange line), and Max Limit (orange line). The KLIPPEL logo is in the top right corner.</p> <p>Note: The histogram cells are defined as right closed (left open) intervals. Only the very first cell is a right closed (left closed) interval to ensure that all single values are taken into account.</p>
<p><b>Time Course</b></p>	<p>Box plot of single values versus time:</p>  <p>The figure shows a box plot of 'fs' values versus 'hour'. The x-axis is labeled 'hour' and ranges from 0 to 25. The y-axis is labeled 'fs' and ranges from 34 to 56. A red line connects the medians of the box plots. Two horizontal orange lines represent the 'Min Limit' and 'Max Limit' at approximately 33 and 55 respectively. The KLIPPEL logo is in the top right corner.</p> <p>Note: In the current version it is not possible to generate time course plots! This feature will come with future updates of QC6. Please use the QC4 version of the YieldStatistics for Time Course analysis. Contact <a href="mailto:support@klippel.de">support@klippel.de</a> for further information.</p>

### 3 Input Parameters

<b>Location of log files</b>	Directory containing the log-files to parse. Log-files may be in sub folders, files are searched recursively. Absolute and relative (to the database's location) paths are allowed. If <i>Folder</i> is empty, the database's location is used.
<b>Name of statistic</b>	Name of test to be shown in HTML output window. If not specified the project's name is extracted from the database's parent folder.
<b>Date Filter</b>	Date range of statistics. <ul style="list-style-type: none"> <li>• YYYYMMDD or YYMMDD: The single date notation specifies the start date of the statistics. The end date is set to now.</li> <li>• YYYYMMDD-YYYYMMDD or YYMMDD-YYMMDD: The date range notation specifies the start date and the end date of the statistics. Use either 2-digit or 4-digit year notation!</li> </ul> <p>The short notation can be used for flexible and fast date/time range input. It is explained with examples in section <i>Placeholder Tokens for Date/Time Input</i> below. If the short notation is used, the <i>Time Filter</i> will be ignored.</p> <p>Please note, that the date range is displayed in the configuration table in the window <i>HTML out</i>.</p> <p>If <i>Date Filter</i> is empty, all available DUTs are considered, until now.</p>
<b>Time Filter</b>	Time range of statistics. This parameter specifies the exact time at the beginning and end of the date range. <b>HH:MM:SS:DS or HH:MM:DD:DS-HH:MM:SS:DS</b> The time range is not interpreted if the short notation for date is used. The time notation can be a range or a single time specifier. If not specified as range the end time is set to 23:59:59:9. The time specifier can be coarse (e.g. HH:MM, HH:MM:SS, HH-HH, HH:MM:SS-HH:MM:SS) and asymmetric (e.g. HH-HH:MM:SS). Please note, that the time range is also displayed in the configuration table in window <i>HTML out</i> .
<b>Operator</b>	Consider only specified operator. You could also separate multiple operators by line feed. This parameter is case-insensitive.
<b>define SN prefix</b>	Check this checkbox if you want to explicitly define a prefix for filtering the serial numbers. If not checked, any prefix of the serial numbers is ignored.
<b>SN prefix</b>	Define your serial number prefix here (Example: "ABC_"). If Prefix is empty only serial numbers without any prefix are shown. Note: Prefixes and serial numbers are separated at the last non-digit (e.g. whole SN = "g7h_38" → prefix = "g7h_"; number = "38")
<b>SN start value</b>	Start value of range of serial numbers to be filtered.
<b>SN end value</b>	End value of range of serial numbers to be filtered.
<b>Ignore SV</b>	List of single values that should not be considered in the verdict and single value statistics. These single values are not exported to the CSV file, too. This parameter is case sensitive. Multiple single value names could be separated by line feed.
<b>CSV-File</b>	Writes all parsed data into a merged CSV file. The file is located in the specified folder. Absolute and relative (to the database's location) paths are allowed. Note: If specified file exists already, it will be overwritten without any warning.

<p><b>Edit plots</b></p>	<p>Configuration for time course or distribution analysis of single values. The configuration is done in a table with each row representing a separate plot. Single rows could easily be commented by using “//” at the beginning of the parameter <b>SV</b>. There is no limitation on the number of plots. The first line is the heading of the configuration table..</p> <p>Available parameters are:</p> <ul style="list-style-type: none"> <li>• <b>SV</b>: specifies the name of the single value to be analyzed. (Example: ‘Fs’)</li> <li>• <b>Type</b>: specifies if the chart is used for distribution or time course analysis. Valid entries are either <b>dist</b> or <b>time</b>.</li> <li>• <b>Filter</b>: (optional) configures the range of considered single value. This can be a string <b>Outlier</b> (for automatic determination of outliers with GRUBBS’ algorithm), <b>Limits</b> (everything within the QC limit range for this single value) or a range <b>[start,end]</b> (definition of absolute range).</li> <li>• <b>Bins</b>: (optional, only for distribution plots) number of bins for histogram plot. Individual start and stop values for histogram bins may be defined with a vector. This parameter is recessive against the parameter <b>Filter</b>.</li> <li>• <b>Unit</b>: (optional, only for time course analysis) specification of time intervals, possible choices are <b>ds</b> (1/10<sup>th</sup> of a second), <b>second</b>, <b>minute</b>, <b>hour</b>, <b>day</b>, <b>week</b>, <b>month</b>, <b>quarter</b> and <b>year</b>.</li> <li>• <b>Interval</b>: (optional, only for time course analysis) quantity of intervals specified by <b>Unit</b>.</li> </ul> <p>Example:</p> <pre>//SV          'TYPE' 'FILTER'      'BINS' 'UNIT' 'INTERVAL' //Re          'dist' '[1,10]'    '30'   ''    '' //Level       'time' 'outlier'    ''     'min' '15' 'Re           'dist' 'limit' ''     ''     ''</pre>
--------------------------	--

#### 4 Placeholder Tokens for Date/Time Input

<p><b>YYYYMMDD or YMMDD</b></p>	<p>The single date notation specifies the start date of the statistics. The end date is set to current time. Use either 2-digit or 4-digit year notation!</p>
<p><b>YYYYMMDD-YYYYMMDD or YMMDD-YMMDD</b></p>	<p>The date range notation specifies the start date and the end date of the statistics. Use either 2-digit or 4-digit year notation!</p>
<p><b>Short notation of statistic's date/time range</b>  <b>[Y Q M W D H {- n}]</b></p>	<p>Note: Since the short notation defines the time range as well, the time notation is not considered if the short notation is used.</p> <p>The short notation specification can relate to either:</p> <ul style="list-style-type: none"> <li>• Year: <b>Y</b>,</li> <li>• Quarter of year (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Dec): <b>Q</b>,</li> <li>• Month of year: <b>M</b>,</li> <li>• Week of year (starting with Sunday): <b>W</b>,</li> <li>• Day of month: <b>D</b> or</li> <li>• Hour of day: <b>H</b>.</li> </ul> <p>Only one specifier for a period of time can be used:</p> <ul style="list-style-type: none"> <li>• <b>Alone</b> to specify current period of time,</li> <li>• <b>With minus sign</b> to specify last period of time or</li> </ul>

<ul style="list-style-type: none"> <li>• <b>With a positive number</b> to specify an arbitrary period of time</li> </ul> <p>Examples (assuming now is Wednesday 13th July 2011, 11:15:09:2, (2011-07-13 11:15:09:2)):</p>	
Example	Interpreted date/time range
Y	2011-01-01 00:00:00:0 - 2011-07-13 11:15:09:2
Q	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
M	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
W	2011-07-10 00:00:00:0 - 2011-07-13 11:15:09:2
D	2011-07-13 00:00:00:0 - 2011-07-13 11:15:09:2
H	2011-07-13 11:00:00:0 - 2011-07-13 11:15:09:2
Y-	2010-07-13 11:15:09:2 - 2011-07-13 11:15:09:2
Q-	2011-04-13 11:15:09:2 - 2011-07-13 11:15:09:2
M-	2011-06-13 11:15:09:2 - 2011-07-13 11:15:09:2
W-	2011-07-06 11:15:09:2 - 2011-07-13 11:15:09:2
D-	2011-07-12 11:15:09:2 - 2011-07-13 11:15:09:2
H-	2011-07-13 10:15:09:2 - 2011-07-13 11:15:09:2
Y9, Y2009	2009-01-01 00:00:00:0 - 2009-12-31 23:59:59:9
Q4	2010-10-01 00:00:00:0 - 2010-12-31 23:59:59:9
M7	2011-07-01 00:00:00:0 - 2011-07-13 11:15:09:2
W3	2011-01-09 00:00:00:0 - 2011-01-15 23:59:59:9
D31	2011-06-30 00:00:00:0 - 2011-06-30 23:59:59:9
H12	2011-07-12 12:00:00:0 - 2011-07-12 12:59:59:9

## 5 Handling Large Amounts of Data

In general, there is no limitation of the size of data to be analyzed. To avoid trouble with plotting of very huge data sets there is a threshold of 100.000 data sets. If the requested single value has more data sets, it will be quantized to reduce the data. If the number of samples is below this threshold there is no change of the data.

In case of quantization of the single values a relative error of <0.1% is guaranteed.

Find explanations for symbols at:

<http://www.klippel.de/know-how/literature.html>

Last updated: September 30, 2019

